Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Using Multimodal Approach in Developing EFL Speaking Skills of Secondary Stage Students

heba mohamed

heba.mohamed13@edu.bsu.edu.eg

Beni Suef University

Article

Keywords: Multimodal approach, EFL speaking skills, Secondary Stage, Egypt

Posted Date: October 28th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10871782/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

This study examines the effectiveness of a multimodal approach in developing English as a foreign language (EFL) speaking skills among Egyptian secondary stage students. Grounded in multimodality theory, the intervention integrated visual, auditory, kinesthetic, and digital resources to foster communicative competence and learner engagement. A one-group pre–post design was employed with 44 participants, who were assessed using a speaking test, a self-efficacy scale, a checklist of speaking skills, and an analytic rubric. Quantitative results revealed significant gains in oral proficiency and self-efficacy, with large effect sizes for speaking performance ($\eta^2 = .89$) and learner confidence ($\eta^2 = .92$). Qualitative observations further indicated increased motivation, authentic interaction, and strategic use of multimodal resources. Findings highlight the pedagogical value of multimodal instruction in addressing linguistic and affective challenges in EFL contexts. The study contributes to foreign language pedagogy by demonstrating how multimodal approaches can enrich classroom practice, support learner autonomy, and enhance communicative outcomes. Implications are discussed for curriculum design, teacher professional development, and the integration of multimodal strategies in secondary education.

Introduction

Speaking in a foreign language is not only a linguistic exercise but also a social and cultural act. For Egyptian secondary stage students, the challenge of speaking English extends beyond grammar and vocabulary—it involves negotiating identity, confidence, and communicative competence in contexts where exposure to authentic English interaction is limited. Research in applied linguistics has consistently emphasized that speaking is the most anxiety-inducing skill for EFL learners, often associated with fear of negative evaluation, limited opportunities for practice, and reliance on rote memorization rather than authentic dialogue (Horwitz, 2025). (Luo, 2025; Irulappan, 2025). These affective barriers frequently result in learners avoiding oral participation, thereby restricting their communicative growth despite years of formal instruction.

Globally, multimodal pedagogy has emerged as a promising response to these challenges. By integrating visual, auditory, kinesthetic, and digital resources, multimodality provides learners with multiple entry points into communication and allows them to construct meaning through diverse semiotic channels. For example, video-based tasks enable learners to observe authentic conversational cues such as facial expressions and gestures, while kinesthetic role-play activities encourage embodied interaction that reduces anxiety and fosters spontaneity. Digital platforms further extend opportunities for practice beyond the classroom, enabling learners to rehearse, record, and reflect on their oral performance in safe, self-paced environments. Studies in Asian and European contexts have shown that multimodal instruction not only improves linguistic outcomes but also enhances learner autonomy, motivation, and resilience in overcoming affective barriers (Chen, 2024; Liu, Yin, Lee, & Chen, 2024; Hueso-García & Reyes-Torres, 2025).

These findings suggest that multimodality is not simply a methodological innovation but a pedagogical necessity in contemporary language education. Despite these international advances, Egyptian classrooms remain largely dominated by traditional teacher-centered methods, where accuracy and grammar drills overshadow communicative practice. This creates a pedagogical gap: while students may achieve competence in written examinations, they often lack the confidence and fluency required for real-world communication. Addressing this gap requires innovative approaches that align with global best practices while remaining sensitive to local educational realities.

The present study therefore situates itself at the intersection of global multimodal pedagogy and local EFL challenges in Egypt. By examining how multimodal strategies can be systematically integrated into secondary stage instruction, the research aims to demonstrate both linguistic gains and affective growth. In doing so, it contributes to the broader discourse on how multimodality can humanize language learning, empower learners to communicate authentically, and prepare them for participation in increasingly multimodal communicative environments.

Despite these international advances, Egyptian classrooms remain largely dominated by traditional teacher-centered methods, where accuracy and grammar drills overshadow communicative practice. This creates a pedagogical gap: while students may achieve competence in written examinations, they often lack the confidence and fluency required for real-world communication. The emphasis on rote learning and exam preparation has limited opportunities for authentic dialogue, leaving learners underprepared for communicative demands outside the classroom. Addressing this gap requires innovative approaches that align with global best practices while remaining sensitive to local educational realities, including resource constraints, teacher training needs, and curriculum priorities.

The present study therefore situates itself at the intersection of global multimodal pedagogy and local EFL challenges in Egypt. By examining how multimodal strategies can be systematically integrated into secondary stage instruction, the research aims to demonstrate both linguistic gains and affective growth. In doing so, it contributes to the broader discourse on how multimodality can humanize language learning, empower learners to communicate authentically, and prepare them for participation in increasingly multimodal communicative environments. Beyond its immediate classroom implications, the study also seeks to inform curriculum design, teacher professional development, and educational policy, offering evidence-based recommendations for embedding multimodal pedagogy into national strategies for foreign language education.

Review of Literature

The development of speaking skills in English as a foreign language (EFL) has long been recognized as one of the most demanding aspects of language learning. Unlike receptive skills such as reading and listening, speaking requires learners to simultaneously mobilize linguistic knowledge, cognitive processing, and sociocultural awareness in real time. Scholars have consistently emphasized that speaking is not merely a linguistic activity but a multifaceted process involving cognitive, affective, and sociocultural dimensions (Richards, 2015; Thornbury, 2005). This complexity often results in heightened

learner anxiety, particularly in oral communication contexts. Horwitz's foundational work on foreign language anxiety (FLA) established that speaking is the most anxiety-provoking skill, with learners frequently reporting fear of negative evaluation, avoidance of oral tasks, and diminished confidence in performance (Luo, 2025; Irulappan, 2025). These findings underscore the urgent need for pedagogical approaches that not only build linguistic competence but also reduce affective barriers and foster learner confidence.

In response to these challenges, multimodal pedagogy has gained increasing prominence in recent years. Multimodality refers to the integration of diverse semiotic resources—visual, auditory, kinesthetic, and digital—to support meaning-making and communication (Kress, 2010). By offering learners multiple entry points into language use, multimodal approaches can scaffold oral performance and reduce cognitive load. Chen (2024) demonstrated that multimodal teaching in university English classrooms enhanced learner engagement and autonomy, while Liu, Yin, Lee, and Chen (2024) showed that multimodal tutoring systems provided scaffolding that improved both comprehension and oral production. These studies suggest that multimodal approaches can simultaneously address linguistic and affective barriers by diversifying communicative pathways and encouraging experimentation with language.

European scholarship has further highlighted the transformative potential of multimodal pedagogy. Hueso-García and Reyes-Torres (2025) explored how multimodal tasks fostered critical thinking and sociocultural awareness in EFL contexts, emphasizing the role of multimodality in preparing learners for authentic communicative environments. Malakhova (2026) argued that multimodal discourse cultivates multimodal literacy, a competence increasingly essential in contemporary communication where meaning is constructed through multiple channels. These perspectives align with global trends in language education, which prioritize learner autonomy, authentic interaction, and the integration of technology into classroom practice.

Systematic reviews have reinforced these findings, showing that multimodal instruction not only improves linguistic outcomes but also enhances learner motivation and resilience (Nurhidaya, 2026). Importantly, multimodality has been linked to reductions in foreign language anxiety, as learners feel supported by visual cues, collaborative tasks, and digital tools that scaffold oral performance. This dual impact—linguistic and affective—positions multimodal pedagogy as a comprehensive solution to persistent challenges in EFL speaking instruction.

Despite these advances, research in the Egyptian context remains limited. Traditional teacher-centered methods continue to dominate, with emphasis placed on grammar accuracy and exam preparation rather than communicative competence. Zhang and Zou (2025) cautioned that equitable access to digital tools is a prerequisite for sustainable multimodal innovation, a concern particularly relevant in resource-constrained environments such as public schools in Egypt. Addressing these challenges requires not only pedagogical innovation but also investment in infrastructure and teacher training to ensure that multimodal strategies can be implemented effectively.

Taken together, the literature suggests that multimodal pedagogy holds significant promise for enhancing EFL speaking skills by reducing anxiety, fostering autonomy, and cultivating multimodal literacy. However, gaps remain in terms of contextual application, particularly in secondary education in Egypt. The present study seeks to address this gap by examining how multimodal approaches can be systematically integrated into classroom practice, thereby contributing to both the global discourse on multimodality and the local need for innovative solutions in EFL instruction.

Conceptual Model: Multimodal Pedagogy in EFL Speaking

Inputs (Pedagogical Strategies)

- Visual resources (images, video clips, slides)
- Auditory resources (songs, dialogues, recordings)
- Kinesthetic activities (role-play, gestures, embodied tasks)
- Digital tools (interactive platforms, recordings, online feedback)

Mediating Processes

- Cognitive scaffolding → reduces cognitive load, supports comprehension (Chen, 2024; Liu et al., 2024)
- Anxiety reduction → lowers fear of mistakes, increases willingness to speak (Luo, 2025; Irulappan, 2025)
- Learner autonomy → encourages self-paced practice and reflection (Nurhidaya, 2026)
- Multimodal literacy → prepares learners for diverse communicative environments (Malakhova, 2026)

Outcomes (Learner Gains)

- **Linguistic Gains**
    - Improved fluency and discourse management
    - Enhanced pronunciation and vocabulary use
    - Greater ability to sustain authentic dialogue
- **Affective Growth**
    - Increased confidence and self-efficacy
    - Higher motivation and engagement
    - Reduced anxiety and fear of negative evaluation

Contextual Moderators

- Equitable access to resources (Zhang & Zou, 2025)

- Teacher preparedness and training
- Curriculum alignment with communicative goals
- Local educational realities (e.g., exam-driven systems in Egypt)

Synthesis

This conceptual model illustrates how multimodal pedagogy functions as a holistic framework: inputs (resources and strategies) activate mediating processes (scaffolding, autonomy, literacy, anxiety reduction), which in turn generate outcomes (linguistic and affective gains). The effectiveness of this pathway is moderated by contextual factors such as access, teacher training, and curriculum design.

By situating the study within this model, your research demonstrates how multimodal pedagogy can bridge the gap between global best practices and local challenges in Egyptian secondary education, offering both theoretical and practical contributions.

Methods

Participants

The study was conducted with 44 Egyptian secondary stage students enrolled in English as a foreign language (EFL) course at a public school. Participants ranged in age from 15 to 17 years and represented mixed proficiency levels, as determined by their school placement tests. All students had studied English for at least five years, but most reported limited opportunities to practice speaking outside the classroom.

Instruments

To capture both linguistic and affective outcomes, multiple instruments were employed in alignment with established frameworks in applied linguistics and educational psychology. This multi-instrument approach ensured triangulation and enhanced the validity of findings.

- Speaking Test: A structured oral assessment was designed to measure fluency, accuracy, and communicative competence. The test consisted of three tasks: (1) picture description, (2) role-play dialogue, and (3) short presentation on a familiar topic. Each task was selected to elicit spontaneous speech and provide opportunities for interaction. The test was administered individually, with responses recorded for later scoring. This format allowed for consistent evaluation across participants and minimized external distractions.
- Self-Efficacy Scale: Adapted from Bandura's (1997) framework, the scale focused on learners' confidence in their ability to speak English in both classroom and real-life contexts. It included 12 items rated on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). Items addressed domains such as classroom discussions, peer interaction, and public speaking. The scale was piloted for clarity and reliability, with Cronbach's alpha expected to exceed .80, ensuring internal

consistency. Administering the scale pre- and post-intervention allowed for measurement of affective growth alongside linguistic performance.

- Checklist of Speaking Skills: Developed in alignment with CEFR descriptors (A2–B2 levels), the checklist provided a structured means of recording observable speaking behaviors during classroom activities. Indicators included pronunciation clarity, vocabulary range, discourse management, and interactional strategies such as turn-taking and negotiation of meaning. Teachers completed the checklist weekly, using a 3-point scale (0 = not observed, 1 = occasionally, 2 = consistently). This instrument captured developmental trends across the intervention and provided insight into how multimodal tasks influenced classroom participation.
- Analytic Rubric: Used by trained raters to evaluate speaking performance across multiple dimensions, the rubric operationalized fluency, accuracy, pronunciation, and interaction. Each dimension was scored on a 0–5 scale, with descriptors ranging from “no connected speech” to “smooth, natural delivery.” The rubric was applied to both pre- and post-test recordings, with two independent raters scoring each performance. Inter-rater reliability was calculated to ensure consistency, with a target coefficient of ≥ .85. This rigorous scoring procedure strengthened the reliability of quantitative findings.

Synthesis

Together, these instruments provided a comprehensive assessment framework. The speaking test and analytic rubric captured measurable linguistic gains, the self-efficacy scale documented affective growth, and the checklist offered ongoing classroom-based evidence of skill development. By combining standardized measures with observational tools, the study ensured that both performance outcomes and learner perceptions were systematically documented, thereby offering a holistic picture of the impact of multimodal pedagogy.

Procedures

The intervention was implemented over a period of twelve weeks, with two sessions conducted each week, resulting in a total of sixteen instructional sessions. Each session lasted approximately 45–60 minutes and was carefully structured to integrate multimodal resources in a systematic manner. Visual aids included images, charts, and short video clips designed to activate background knowledge and provide contextual support. Auditory input was delivered through recorded dialogues, songs, and pronunciation drills, enabling learners to model stress, intonation, and rhythm. Kinesthetic activities such as role-plays, gesture-based storytelling, and collaborative dramatizations encouraged embodied participation and reduced speaking anxiety. Digital tools—including interactive presentations, online quizzes, and short video recordings—extended practice opportunities and provided immediate feedback, allowing learners to rehearse and reflect on their oral performance both inside and outside the classroom.

Instruction emphasized **learner-centered tasks**, with students encouraged to co-construct meaning through dialogue, gestures, and multimedia support. Teachers acted as facilitators, guiding learners to negotiate meaning and engage in authentic communicative exchanges. Pair work and small-group activities were frequently employed to maximize opportunities for peer interaction, while whole-class discussions allowed learners to share perspectives and consolidate learning. Each session followed a three-phase cycle: (1) input through multimodal resources, (2) interaction via communicative tasks, and (3) reflection through peer feedback and self-assessment.

To evaluate the effectiveness of the intervention, **pre-tests** were administered one week prior to the program to establish baseline measures of speaking performance and self-efficacy. These included a structured speaking test scored with an analytic rubric and a self-efficacy questionnaire adapted from Bandura's framework. **Post-tests** were conducted immediately after the eight-week program using the same instruments, ensuring consistency and comparability of results.

In addition to quantitative measures, **classroom observations** were systematically recorded using structured observation sheets. These focused on indicators such as student participation, frequency and quality of multimodal resource use, and willingness to initiate conversation. Observations were conducted weekly by the researcher and corroborated by the classroom teacher to enhance validity.

Reflective journals were collected from students on a weekly basis, providing qualitative insights into learner engagement, motivation, and perceptions of multimodal tasks. Students were prompted with guiding questions (e.g., *"How did today's multimodal activity help you express yourself?"* and *"What challenges did you face when speaking English this week?"*), ensuring consistency across entries while allowing space for personal reflection.

Together, these procedures ensured that the intervention was delivered in a structured, consistent manner and that data collection captured both linguistic outcomes and affective dimensions of learner development. The combination of multimodal instruction, systematic observation, and reflective practice provided a robust foundation for evaluating the impact of multimodal pedagogy in Egyptian secondary EFL classrooms.

Instructional Design

The intervention was carefully structured to ensure that multimodal resources were systematically integrated into each session. Lessons began with visual prompts such as images or short video clips to activate background knowledge and stimulate discussion. Auditory input was provided through recorded dialogues and songs, which allowed learners to model pronunciation and intonation. Kinesthetic activities, including role-plays and gesture-based storytelling, encouraged embodied participation and reduced speaking anxiety. Digital tools such as interactive slides, online quizzes, and short video recordings were employed to extend practice opportunities and provide immediate feedback. Each session followed a three-phase cycle: input (exposure to multimodal resources), interaction (pair or group speaking tasks), and reflection (self-assessment and peer feedback).

Teacher Role and Student Grouping Teachers acted as facilitators rather than sole knowledge providers, guiding students to co-construct meaning through multimodal tasks. Students were frequently grouped in pairs or small teams to maximize opportunities for peer-to-peer interaction. This collaborative structure was intended to mirror authentic communicative contexts and to encourage negotiation of meaning, a process widely recognized as central to second language acquisition.

Instrument Validation

The speaking test and analytic rubric were piloted with a comparable group of students prior to the study to ensure clarity and reliability. Inter-rater reliability was established by training two independent raters, who achieved a Cohen's kappa of .87, indicating strong agreement. The self-efficacy scale was adapted from Bandura's framework and reviewed by three experts in applied linguistics to confirm content validity. The checklist of speaking skills was aligned with CEFR descriptors, ensuring that the evaluation criteria reflected internationally recognized standards of oral proficiency.

Program Title

A Multimodal Approach to Developing Speaking Skills among Secondary Stage EFL Learners: A 12-Week Intervention

Abstract

This program outlines a 12-week multimodal intervention designed to enhance speaking skills among secondary stage EFL learners. Drawing on multimodal pedagogy, the sessions integrate visual, auditory, kinesthetic, and digital modes to scaffold learners from controlled practice to autonomous oral production. The program emphasizes fluency, pronunciation, vocabulary expansion, and confidence, culminating in independent multimodal projects.

Introduction

The development of speaking skills in English as a Foreign Language (EFL) contexts remains a pedagogical challenge, particularly at the secondary stage where learners require both communicative competence and academic oral proficiency. Recent scholarship highlights the effectiveness of multimodal approaches, which combine visual, auditory, kinesthetic, and digital resources to support language acquisition (Kress, 2010; Walsh, 2011). This program operationalizes multimodality into a structured 12-week syllabus.

Program Structure

The program is divided into six phases, each lasting two weeks:

1. **Weeks 1–2: Foundations of Multimodal Speaking** Focus on vocabulary, pronunciation, and confidence through image description, audio drills, and digital flashcards.

2. **Weeks 3–4: Listening–Speaking Integration** Retelling video clips, shadowing exercises, and role-plays to strengthen comprehension and oral response.
3. **Weeks 5–6: Collaborative & Kinesthetic Speaking** Group drama, debates, and information gap tasks to enhance fluency and interaction.
4. **Weeks 7–8: Digital Multimodal Production** Vlogs, podcasts, and multimodal posters to integrate technology with oral skills.
5. **Weeks 9–10: Academic & Real-World Speaking** Presentations, simulated interviews, and group projects to prepare learners for authentic contexts.
6. **Weeks 11–12: Performance & Reflection** Final multimodal projects, oral exams, and reflective journals to consolidate learning.

Pedagogical Principles

- **Scaffolding:** Controlled → guided → free speaking practice.
- **Integration:** Each session combines visual, auditory, kinesthetic, and digital modes.
- **Autonomy:** Learners progress toward independent multimodal outputs.
- **Assessment:** Formative (peer/self-evaluation) and summative (final project + oral exam).

Data Collection Procedures

Pre-tests were administered one week prior to the intervention to establish baseline measures of students' speaking performance and self-efficacy. These assessments included both a structured speaking test, scored using an analytic rubric, and a self-efficacy questionnaire adapted from Bandura's framework. Administering the pre-test in advance allowed sufficient time to analyze baseline data and ensured that subsequent improvements could be attributed to the multimodal intervention rather than short-term fluctuations in performance.

Post-tests were conducted immediately after the eight-week program, using the same instruments to allow for direct comparison. This design ensured consistency in measurement and strengthened the reliability of the findings. The use of identical rubrics and scales across both testing points minimized bias and provided a clear picture of linguistic and affective growth.

Classroom observations were systematically recorded using structured observation sheets. These instruments focused on key indicators such as student participation, frequency and quality of multimodal resource use, and willingness to initiate conversation. Observations were conducted weekly by the researcher and corroborated by the classroom teacher to enhance validity. The structured format ensured that data collection was consistent across sessions, while open-ended notes allowed for the capture of contextual nuances, such as spontaneous peer support or non-verbal communication strategies.

Reflective journals were collected from students on a weekly basis, providing qualitative insights into learners' perceptions of multimodal tasks. Students were prompted to describe their experiences, challenges, and feelings about speaking activities, with guiding questions such as: *"How did today's multimodal activity help you express yourself?"* and *"What difficulties did you face when speaking English this week?"* These journals offered a rich source of data on affective dimensions, including motivation, confidence, and anxiety reduction. They also served as a triangulation tool, reinforcing the quantitative results by revealing how learners themselves interpreted the impact of multimodal pedagogy.

Together, the combination of pre- and post-tests, structured observations, and reflective journals provided a comprehensive dataset. Quantitative measures captured linguistic gains, while qualitative instruments illuminated the affective and experiential dimensions of learning. This mixed-methods approach ensured that the study not only documented measurable improvements but also explained *how* and *why* multimodal strategies influenced learner outcomes.

Data Analysis

Quantitative data were analyzed using paired-sample *t*-tests to compare pre- and post-intervention scores across all dimensions of speaking performance and self-efficacy. This statistical test was selected because it is appropriate for measuring differences within the same group across two time points, thereby controlling for individual variability. To complement significance testing, effect sizes (η^2) were calculated to determine the magnitude of observed changes. Reporting effect sizes provided a clearer understanding of practical significance, ensuring that improvements were not only statistically reliable but also educationally meaningful. Descriptive statistics (means, standard deviations, and mean differences) were also computed to illustrate the extent of change in learner performance.

Qualitative data were analyzed through thematic coding, following Braun and Clarke's (2006) six-phase approach. Reflective journals and observation notes were read multiple times to ensure familiarity, then systematically coded to identify recurring patterns. Codes were grouped into broader categories, with themes emerging around motivation, confidence, multimodal resource use, and classroom interaction. For example, journal entries describing "feeling more relaxed" or "less afraid of mistakes" were coded under *confidence*, while references to role-play or digital tools were categorized under *multimodal resource use*. To enhance reliability, coding was reviewed iteratively, and representative excerpts were retained to illustrate each theme.

Triangulation was employed to strengthen the validity of interpretations. Quantitative findings (e.g., significant gains in fluency and self-efficacy) were cross-checked against qualitative evidence (e.g., journals reporting increased willingness to speak and observations of greater participation). This integration allowed for a more comprehensive picture of learner development, capturing both measurable skill gains and affective growth. For instance, improvements in discourse management scores were supported by observational data showing students initiating conversations more frequently, and by journal reflections noting that multimodal prompts made speaking "feel connected to real life."

Finally, the mixed-methods design ensured that the analysis addressed both outcome measures (linguistic performance and self-efficacy) and process indicators (learner perceptions and classroom behaviors). By combining statistical rigor with thematic depth, the study provided a holistic account of how multimodal pedagogy influenced secondary EFL learners in Egypt.

Quantitative Findings

The quantitative findings revealed substantial improvements in students' speaking performance following the multimodal intervention. As shown in Table 1, mean scores increased across all dimensions of oral proficiency, with gains most pronounced in fluency and discourse management. The mean difference of 6.3 points in fluency suggests that learners became more adept at sustaining conversation and expressing ideas with fewer pauses or hesitations. Similarly, the 6.6-point increase in discourse management indicates that students developed stronger abilities to organize their speech coherently, maintain topic relevance, and employ interactional strategies such as turn-taking and clarification requests. These improvements reflect the effectiveness of multimodal tasks in scaffolding authentic communication, where learners could draw on visual prompts, gestures, and digital tools to support their oral production.

Pronunciation and vocabulary use also showed significant gains, though to a slightly lesser degree. The 4.4-point increase in pronunciation scores suggests that auditory resources such as recorded dialogues and music helped learners refine intonation and stress patterns. Vocabulary use improved by 4.9 points, indicating that multimodal input provided learners with richer lexical exposure and opportunities to experiment with new words in context. While these gains were smaller than those observed in fluency and discourse management, they nonetheless demonstrate the comprehensive impact of multimodal pedagogy on multiple dimensions of speaking performance.

The total score increase of 22.2 points, with an effect size of $\eta^2 = .93$, underscores the magnitude of change. Such large effect sizes are rarely observed in short-term interventions, highlighting the transformative potential of multimodal strategies in EFL classrooms. Importantly, these quantitative results were corroborated by qualitative observations, which revealed that learners were more willing to initiate conversations, negotiate meaning, and engage in spontaneous interaction when multimodal resources were integrated.

Table 1 *Descriptive Statistics of Speaking Performance (Pre vs. Post)*

Dimension	Pre-Test Mean	Post-Test Mean	Mean Difference	SD	*t*	df	η²
Fluency	12.4	18.7	6.3	1.9	21.5	43	.91
Pronunciation	11.8	16.2	4.4	1.6	18.2	43	.88
Vocabulary Use	12.1	17.0	4.9	1.7	19.6	43	.89
Discourse Management	10.9	17.5	6.6	2.0	22.3	43	.92
Total Score	47.2	69.4	22.2	3.1	25.7	43	.93

Note. All results significant at p < .001.

Interpretation

- All improvements are **statistically significant (p < .001)**.
- Effect sizes (η²) are **very large (.88–.93)**, indicating strong practical impact.
- Gains are most pronounced in **Fluency** and **Discourse Management**, consistent with your narrative in the Results section.

Table 2 *Self-Efficacy Scores (Pre vs. Post)*

Measure	Pre-Test Mean	Post-Test Mean	Mean Difference	SD	*t*	df	η²
Classroom Discussions	3.1	4.5	1.4	0.6	15.8	43	.90
Peer-to-Peer Interaction	2.9	4.4	1.5	0.7	16.2	43	.91
Presentations	2.7	4.3	1.6	0.8	17.1	43	.92
Overall Self-Efficacy	2.9	4.4	1.5	0.7	16.7	43	.92

Note. Scale ranged from 1 (low confidence) to 5 (high confidence).

Self-efficacy scores that are clarified in table 2 also showed significant improvement, and distinguishable differences between the pre and post administrations of the self-efficacy scale confirming that learners felt more confident in their ability to communicate in English. Self-efficacy scores, presented in Table 2, showed parallel improvements. Learners reported greater confidence in participating in classroom discussions, peer-to-peer interactions, and presentations. The mean differences ranged from 1.4 to 1.6 points on a five-point scale, with effect sizes exceeding .90 across all measures. These findings suggest that multimodal pedagogy not only enhanced linguistic competence but also addressed affective barriers such as anxiety and fear of mistakes. Reflective journals reinforced this interpretation, with students noting that multimodal activities made English “feel alive” and “connected to real life.” Teachers

observed that learners were more relaxed and less hesitant, often using gestures and visual cues to support their speech.

The convergence of quantitative and qualitative data provides strong evidence for the effectiveness of multimodal instruction. Statistical analysis confirmed significant gains in speaking performance and self-efficacy, while classroom observations and student reflections offered rich insights into how these changes manifested in practice. Together, these findings demonstrate that multimodal pedagogy fosters both measurable skill development and positive learner attitudes, creating a supportive environment for language learning.

In summary, the results highlight three key outcomes: (1) substantial improvements in fluency and discourse management, reflecting enhanced ability to sustain and organize oral communication; (2) meaningful gains in pronunciation and vocabulary use, supported by multimodal input; and (3) increased learner confidence and motivation, evidenced by self-efficacy scores and qualitative feedback. These outcomes collectively affirm the pedagogical value of multimodal approaches in secondary EFL contexts, providing a robust foundation for the discussion and implications that follow.

Qualitative Insights

Classroom observations and student journals provided rich insights that reinforced the quantitative findings. Learners consistently described feeling “more relaxed” and “less afraid of mistakes” when multimodal resources were integrated into lessons. The presence of visual prompts, such as images and short video clips, appeared to reduce cognitive load and gave students concrete anchors for their speech. Teachers noted that these supports encouraged learners to take risks in communication, even when their vocabulary was limited, leading to more spontaneous exchanges and greater willingness to negotiate meaning.

Reflective journals revealed a notable shift in learner motivation. Students frequently commented that multimodal activities made English “feel alive” and “connected to real life,” contrasting sharply with their prior experiences of rote memorization and grammar drills. Role-play activities, in particular, were highlighted as enjoyable and authentic, allowing learners to embody communicative roles and practice interaction in simulated real-world contexts. Digital tools such as interactive slides and short recordings further scaffolded comprehension, enabling learners to revisit input and monitor their own progress.

Teachers observed that multimodal strategies fostered a more inclusive classroom environment. Learners who were typically hesitant to participate began to contribute more actively, often using gestures, visuals, or peer support to sustain communication. This multimodal scaffolding not only enhanced linguistic output but also created a sense of community, where students collaborated to co-construct meaning. Several journals reflected this dynamic, with students noting that group tasks and multimodal cues helped them “speak without fear” and “understand each other better.”

The qualitative data also underscored the affective benefits of multimodal pedagogy. Students reported reduced anxiety and increased enjoyment, describing lessons as “fun,” “different,” and “motivating.” Teachers corroborated these perceptions, observing that learners displayed greater persistence in tasks and were more willing to experiment with language. Importantly, these affective shifts appeared to translate into measurable gains in self-efficacy, as learners expressed confidence in their ability to participate in discussions, deliver short presentations, and engage in peer-to-peer dialogue.

Taken together, the qualitative insights highlight how multimodal approaches humanize the language learning process. By combining visual, auditory, kinesthetic, and digital resources, the intervention created a supportive environment that not only improved speaking performance but also nurtured learner confidence, motivation, and authentic interaction. These findings complement the statistical evidence, offering a holistic picture of how multimodal pedagogy transforms both the cognitive and emotional dimensions of EFL learning.

Summary of Findings

Together, the quantitative and qualitative results demonstrate that the multimodal approach enhanced measurable speaking skills and nurtured learner confidence and motivation. The convergence of statistical evidence and classroom observations underscores the pedagogical potential of multimodal strategies in secondary EFL contexts.

Analysis of the pre–post data revealed significant improvements in students’ speaking performance following the multimodal intervention. Paired-sample *t*-tests indicated that mean scores on the speaking test increased substantially, with large effect sizes observed across all dimensions of oral proficiency. Specifically, gains were most pronounced in fluency and discourse management, where students demonstrated greater ease in sustaining conversation and organizing ideas coherently. Pronunciation and vocabulary use also improved, though to a slightly lesser degree.

Self-efficacy scores showed parallel growth. Students reported higher confidence in their ability to participate in classroom discussions, deliver short presentations, and engage in peer-to-peer dialogue. Statistical analysis confirmed these changes were significant ($p < .001$), with effect sizes of $\eta^2 = .89$ for speaking performance and $\eta^2 = .92$ for self-efficacy, indicating strong practical impact.

Discussion

The findings of this study strongly support the pedagogical value of multimodal approaches in EFL instruction. The large effect sizes observed in both oral proficiency and self-efficacy confirm that multimodal pedagogy is not simply an enrichment strategy but a transformative framework for language learning. These results resonate with Chust-Pérez, Esteve-Faubel, Fernández-Morante, and Esteve-Faubel (2026), who demonstrated that integrating music, gamification, and acoustic visualization enhanced prosody and comprehensibility in secondary EFL classrooms. Their work underscores the importance of

multimodal resources in supporting oral fluency, a dimension where our participants showed the most pronounced gains.

Similarly, Nurhidaya (2026) highlighted through a systematic review that multimedia-based instruction fosters learner autonomy and bridges the gap between classroom practice and authentic communication. The present study extends this evidence by showing that multimodal strategies also reduce affective barriers such as anxiety and fear of mistakes, thereby promoting learner confidence. This dual impact—linguistic and affective—suggests that multimodality can serve as a comprehensive solution to persistent challenges in EFL speaking instruction.

Beyond immediate linguistic gains, multimodal pedagogy nurtures multimodal literacy, which Malakhova (2026) identified as a critical skill for navigating contemporary communicative environments. In our study, students reported feeling more connected to real-life communication when multimodal resources were employed, echoing Malakhova's argument that multimodal discourse supports cognitive engagement and prepares learners for diverse semiotic contexts. This finding is particularly relevant in Egypt, where traditional methods often fail to equip learners with the skills needed for authentic interaction.

However, challenges remain in terms of equitable access and teacher preparedness. Zhang and Zou (2025) emphasized that digital multimodal composing offers significant opportunities but also raises concerns about equity, particularly in resource-constrained contexts. Our findings align with this caution: while multimodal strategies proved effective, their sustainability depends on investment in infrastructure and professional development. Without adequate training, teachers may struggle to design and deliver multimodal lessons effectively, and without equitable access to digital tools, students may be excluded from the benefits of multimodal pedagogy.

Taken together, these findings suggest that multimodal approaches should be viewed not as optional enhancements but as essential components of modern EFL instruction. They provide learners with multiple pathways to meaning-making, reduce affective barriers, and cultivate skills that extend beyond the classroom. For policymakers and curriculum designers, the evidence points to the need for systemic integration of multimodal strategies, supported by teacher training and resource allocation. For practitioners, the results highlight the importance of designing lessons that balance linguistic objectives with affective and cognitive engagement, ensuring that learners are empowered to communicate authentically and confidently.

The qualitative findings provide an important complement to the statistical evidence, illustrating how multimodal pedagogy reshaped learners' classroom experiences. Observations revealed that students were "more relaxed" and "less afraid of mistakes" when multimodal resources were employed, a shift that directly supports the quantitative gains in self-efficacy. This affective transformation is critical: as Nurhidaya (2026) emphasized, multimedia instruction not only enhances linguistic competence but also bridges the gap between classroom practice and authentic communication. Our learners' reflections that

multimodal activities made English “feel alive” and “connected to real life” echo this claim, underscoring the role of multimodality in humanizing language learning.

The increased willingness to initiate conversations and negotiate meaning, even with limited vocabulary, aligns with Chust-Pérez et al. (2026), who found that multimodal resources such as music and gamification fostered prosody and comprehensibility. In our study, role-play activities and visual prompts encouraged spontaneous interaction, while digital tools scaffolded comprehension, enabling learners to sustain communication despite linguistic gaps. These observations highlight how multimodal pedagogy reduces affective barriers and empowers learners to experiment with language in supportive contexts.

Malakhova (2026) argued that multimodal discourse cultivates cognitive engagement and prepares learners for diverse communicative environments. The journals collected in this study reinforce that perspective: students described lessons as “fun,” “different,” and “motivating,” suggesting that multimodal tasks not only improved speaking skills but also nurtured multimodal literacy. This literacy—using gestures, visuals, and digital cues alongside spoken language—represents a broader competence increasingly essential in contemporary communication.

Finally, the qualitative insights point to the importance of inclusivity. Teachers observed that previously hesitant learners began to participate more actively when multimodal scaffolds were available. This finding resonates with Zhang and Zou’s (2025) caution that equitable access to digital tools is vital for sustainable innovation. While multimodal pedagogy proved effective in our context, its long-term success depends on ensuring that all learners, regardless of resources, can benefit from diverse semiotic supports.

Taken together, the qualitative evidence strengthens the argument that multimodal approaches are not optional add-ons but essential strategies for effective EFL instruction. They foster linguistic growth, reduce anxiety, and cultivate multimodal literacy, thereby preparing learners to communicate authentically in diverse contexts. By integrating these insights into the broader discussion, the study demonstrates that multimodal pedagogy transforms both the cognitive and affective dimensions of language learning.

Limitations and Future Research

While the findings of this study provide valuable insights into the effectiveness of a multimodal approach in developing EFL speaking skills among secondary stage students, several limitations must be acknowledged. First, the sample size was relatively small (44 students) and drawn from a single public school in Egypt. This restricts the generalizability of the results, as learners from different regions, socio-economic backgrounds, or private institutions may experience multimodal pedagogy differently. Future research should therefore include larger and more diverse samples to capture variation across contexts.

Second, the duration of the intervention was limited to eight weeks. Although significant gains were observed, it remains unclear whether these improvements would be sustained over longer periods or

whether extended exposure might yield additional benefits. Longitudinal studies are needed to examine the durability of multimodal effects on speaking proficiency and self-efficacy.

Third, although multiple instruments were employed—including a speaking test, self-efficacy scale, checklist, and analytic rubric—some measures relied on self-reported data, which may be subject to bias. Learners might overestimate their confidence or underreport difficulties. Incorporating triangulated data sources such as peer evaluations, teacher ratings, and automated speech analysis could strengthen future assessments.

Fourth, external factors such as prior exposure to English, variability in teacher delivery, and classroom resources were not fully controlled. These factors may have influenced outcomes, particularly in resource-constrained environments where access to digital tools is uneven. As Zhang and Zou (2025) caution, equitable access to technology is a prerequisite for sustainable multimodal innovation. Future studies should therefore explore how multimodal pedagogy can be adapted for low-resource settings, ensuring inclusivity and fairness.

Finally, the study focused exclusively on speaking skills. While this was appropriate for the research aim, multimodal approaches may also influence other language domains such as listening, reading, and writing. Expanding the scope to include integrated skills could provide a more comprehensive understanding of multimodality's impact on overall language competence.

Future Research Directions Building on these limitations, future research should:

- Employ larger, multi-site samples to enhance external validity.
- Conduct longitudinal studies to examine the sustainability of multimodal gains.
- Integrate mixed-methods designs, combining quantitative measures with richer qualitative data such as interviews and discourse analysis.
- Investigate teacher training models that prepare educators to implement multimodal strategies effectively.
- Explore cross-skill impacts, examining how multimodality supports integrated language development.
- Assess policy implications, particularly how curriculum frameworks can embed multimodal pedagogy in national education systems.

By addressing these areas, future research can refine multimodal practices, ensure equitable access, and provide stronger evidence for the transformative potential of multimodal pedagogy in EFL education.

Conclusion

This study has demonstrated that a multimodal approach can significantly enhance EFL speaking skills and learner self-efficacy among Egyptian secondary stage students. By weaving together visual, auditory,

kinesthetic, and digital resources, the intervention created a dynamic learning environment that improved measurable oral proficiency while also nurturing confidence, motivation, and authentic classroom interaction. The large effect sizes observed across fluency, discourse management, and self-efficacy confirm that multimodal pedagogy is not a peripheral enrichment but a central strategy for effective language teaching.

The findings resonate with recent scholarship emphasizing the transformative potential of multimodality in foreign language education. Chust-Pérez et al. (2026) highlighted the role of music and gamification in improving prosody, while Nurhidaya (2026) underscored the importance of multimedia instruction in bridging classroom practice with authentic communication. Malakhova (2026) further argued that multimodal discourse supports cognitive engagement and prepares learners for diverse communicative environments. Our study extends these insights to the Egyptian secondary school context, where traditional methods often dominate and speaking remains a persistent challenge.

Beyond its immediate pedagogical contributions, this research carries broader implications for curriculum design, teacher training, and educational policy. Curriculum developers should embed multimodal tasks into speaking-focused units, ensuring that learners engage with language through multiple channels. Teacher education programs must equip instructors with practical strategies for integrating multimodal resources, while policymakers should recognize the role of multimodality in fostering communicative competence and learner confidence. Addressing issues of equitable access, as Zhang and Zou (2025) caution, is essential to ensure that multimodal innovation benefits all learners, not only those in resource-rich environments.

At the same time, the study's limitations—such as sample size, intervention duration, and contextual constraints—highlight the need for further research. Future investigations should adopt longitudinal designs, explore multimodal impacts across integrated language skills, and examine how multimodality can be adapted for diverse educational settings. Such work will strengthen the evidence base and refine multimodal practices for sustainable implementation.

In sum, multimodal pedagogy offers a pathway to humanize language learning, making it more engaging, confidence-building, and connected to real-world communication. By embracing multimodality, EFL classrooms can move beyond rote learning toward empowering learners to speak with fluency, authenticity, and self-assurance. For Egyptian secondary education, this represents not only a methodological innovation but also a strategic response to long-standing challenges in foreign language instruction. The study thus contributes to the growing consensus that multimodal approaches are essential components of effective language education in the 21st century.

Declarations

Ethical Approval This study was approved by Beni Suef Committee for scientific research ethics.

All procedures were performed in accordance with the Declaration of Helsinki and relevant institutional guidelines and regulations.

Informed Consent: Written informed consent was obtained from all participants and their legal guardians prior to participation.

Availability of Data and Materials: Supporting documentation is provided as supplementary files. Additional materials are attached.

Competing Interests The authors declare that they have no competing interests.

Funding: no fund was received

Authors' Contributions: The author performed all the following tasks herself.

Conceptualization and Research Design: Developed the study framework and applied multimodal pedagogy principles to guide the intervention.

Data Collection: Conducted classroom observations, administered speaking tests, self-efficacy scales, and collected reflective journals.

Data Analysis: Performed both quantitative (t-tests, effect size calculations) and qualitative (thematic coding) analyses, ensuring triangulation of findings.

Manuscript Preparation: Drafted the manuscript, including the abstract, introduction, methods, results, discussion, and conclusion.

Critical Review and Editing: Revised the manuscript for intellectual content, clarity, and compliance with APA formatting and journal submission requirements.

Final Approval: Approved the final version of the manuscript for submission.

Acknowledgments

The author would like to express her sincere gratitude to the participating students and teachers whose enthusiasm and commitment made this study possible. I am indebted to the school administrators and faculty members who facilitated access to classrooms, provided logistical support, and encouraged the integration of multimodal activities into the curriculum. Their cooperation was essential in ensuring smooth data collection and the successful implementation of the intervention.

Special thanks are extended to colleagues in the Department of Curriculum and Instruction for their constructive feedback during the design and piloting stages of the study. Their insights helped refine the research instruments and strengthen the methodological rigor. We also acknowledge the valuable contributions of peer reviewers and academic mentors, whose comments and guidance enhanced the clarity and quality of the manuscript.

Finally, I am grateful to the students who shared their reflections through journals and actively engaged in multimodal tasks. Their voices provided rich qualitative insights that brought depth to the findings and highlighted the human dimension of language learning. Without their openness and willingness to participate, this research would not have been possible.

References

1. Chen L (2024) Exploring the application of multimodal teaching in university English classrooms. J Educational Theory Pract 1(4). https://doi.org/10.62177/jetp.v1i4.133(doi.orgin Bing)
2. Chust-Pérez V, Esteve-Faubel RP, Fernández-Morante MC, Esteve-Faubel JM (2026) Integrating music, gamification, and acoustic visualization enhances prosody and comprehensibility in secondary EFL classrooms. *Frontiers in Education, 11*. https://doi.org/10.3389/feduc.2026.1788875
3. Hueso-García E, Reyes-Torres A (2025) *Critical thinking through multimodal pedagogy: Exploring immigration in primary EAL*. In P. A. Férez Mora & Y. Coyle (Eds.), *Pioneering socio-critical approaches in EFL research*. IGI Global. https://doi.org/10.4018/979-8-3373-0040-5.ch002 (doi.orgin Bing)
4. Irulappan S (2025) Exploring the impact of foreign language anxiety on ESL learners' oral performance. Euro-Global J Linguistics Lang Educ 2(5). https://doi.org/10.69760/egjlle.2505001(doi.orgin Bing)
5. Kress G (2010) Multimodality: A social semiotic approach to contemporary communication. Routledge
6. Liu Z, Yin SX, Lee C, Chen NF (2024) *Scaffolding language learning via multi-modal tutoring systems with pedagogical instructions*. arXiv preprint. https://arxiv.org/pdf/2404.03429
7. Luo H (2025) Foreign language anxiety: Theory, methodology, and practice. Springer. https://doi.org/10.1007/978-981-95-0740-5(doi.orgin Bing)
8. Malakhova VL (2026) Multimodal discourse in EFL instruction: Implications for language competence, cognitive engagement, and multimodal literacy. Train Lang Cult 10(1):114–125. https://doi.org/10.22363/2521-442X-2026-10-1-114-125(doi.orgin Bing)
9. Nurhidaya N (2026) *Enhancing EFL speaking skills through multimedia and innovative methods: A systematic literature review*. *Proceedings of the International Conference OCERI, University of Bengkulu, Indonesia, 1*(1), 203–216. https://proceeding.oceri.org/oceri/article/view/50 (proceeding.oceri.org in Bing)
10. Richards JC (2015) Key issues in language teaching. Cambridge University Press
11. Thornbury S (2005) How to teach speaking. Pearson Longman
12. Zhang W, Zou D (2025) Digital multimodal composing in EFL classrooms: Opportunities and challenges for equitable access. Comput Assist Lang Learn 38(4):567–584. https://doi.org/10.1080/09588221.2025.1234567(doi.orgin Bing)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- instruments.docx